

DOCUMENT RESUME

ED 154 023

TM 007 036

AUTHOR Froman, Robin D.; Cwer, Steven V.
TITLE An Investigation of the Longitudinal Validity of WISC and Bender Scores for Primary School Special Education Students.
PUB DATE Aug 77
NOTE 12p.; Paper presented at the Annual Meeting of the American Psychological Association (85th, San Francisco, California, August 26-30, 1977)
EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; Cognitive Processes; Data Analysis; Diagnostic Tests; Educational Diagnosis; Elementary Education; Elementary School Mathematics; *Identification Tests; *Intelligence Quotient; *Learning Disabilities; Longitudinal Studies; *Perceptual Motor Coordination; *Predictive Validity; Reading Achievement; Special Education; Test Validity
IDENTIFIERS *Bender Visual Motor Gestalt Test; *Wechsler Intelligence Scale for Children

ABSTRACT

The long term predictive validity of the Wechsler Intelligence Scale for Children (WISC) and the Bender Visual Motor Gestalt Test for Children, when used for the educational placement of exceptional students, was investigated. In 1971, 225 learning disabled elementary school students were tested on the WISC and the Bender. The students were between six and seven years old at the time of the first testing. Group IQ, reading and math grade equivalent scores, and educational placement data were collected on 104 of the subjects in 1976. WISC and Bender scores were regressed against math, reading and group IQ data. A discriminant function analysis was conducted to determine the usefulness of the original test score data in predicting present educational placement of the exceptional students. The results of the analyses were mixed, and suggest that caution should be taken in the use of early test score data to make educational decisions regarding the placement of individual students. The usefulness of the early test scores in making long range educational plans and administrative decisions is discussed.
(Author/BW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Steven V.
Owen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

An Investigation of the Longitudinal
Validity of WISC and Bender Scores for Primary
School Special Education Students

by

Robin D. Froman

Steven V. Owen

University of Connecticut

A paper presented at the annual meeting of the American Psychological Association, San Francisco, August, 1977.

ED154023

IM007 036

An Investigation of the Longitudinal
Validity of WISC and Bender Scores for Primary
School Special Education Students

Introduction and Statement of the Problem

Educators and researchers have long been concerned about the early identification of the learning disabled (LD) child. Many methods of identification of the child have been explored in the past (Clements, 1966). Among the more frequently used diagnostic tools are the Bender Visual-Motor Gestalt Test (Bender, 1938) and the Wechsler Intelligence Scale for Children (Wechsler, 1949).

The Bender was originally developed to assess overall visual motor integration skills. It was designed to reflect maturation level and disturbances in visual-motor functioning. The use of the Bender has, however, been extended from its original purpose. It has been employed to assess intellectual level, mental retardation, and brain damage, and to predict school achievement (Bender, 1970). Schulberg and Tolor (1961) report that the Bender is frequently used to identify LD children when questions or organicity are raised. The researchers reported that in a survey of practicing clinicians, eighty percent of the respondents perceived the Bender as having "some" or "great" diagnostic value in identifying organic problems. Graham and Berman (1961), however, must be referred to in order to qualify and moderate the clinicians' faith in the usefulness of the Bender. They point out that while there had been extensive research involving adults and older children and the use of the Bender by the 1960's, relatively little research had been conducted involving younger children.

During the 1960's, however, the use of the Bender in the diagnosis of LD children received considerable attention. A large amount of evidence was reported to support the use of the Bender to discriminate groups of children with organic impairments from those children who were not impaired (McConnell, 1967; Wagner and Murray, 1969; Wiener, 1966). Hanvik, Nelson, Hanson, Anderson, Dressler, and Zarling (1961) reported that scores on the Bender were the single best predictor of organicity in young children when an abnormal EEG was used as the criterion. Clements and Peters (1962) suggested that the Bender was a valuable measure of perception and visual motor coordination, and that the lack of its use in psychodiagnostic examinations may lead to misdiagnosis.

The second test frequently used to identify the LD child, the Wechsler Intelligence Scale for Children (WISC), was developed to test cognitive functioning. Since the collection of the data presented in this paper, the WISC has been revised (WISC-R, Wechsler, 1974). The first version of the WISC was used in this study, so only that research pertaining to the validity and use of the original version will be discussed.

Some investigators claim to have determined a link between learning disorders or minimal brain dysfunctions (MBD) and WISC scores (Clements and Peters, 1962; Eaves, Kendall, and Crichton, 1972; Holroyd, 1968). A review of much of the research exploring the relationship between suspected cerebral impairments as a source of learning disabilities in children and test performance on the WISC is presented by Rourke (1975). He describes two ways in which the WISC has been used as a diagnostic tool in the determination of lateralized motor deficits and in the differential score approach. The differential score approach, which uses a comparison of the Verbal and Performance IQ scores, has been found to be one of the most powerful discriminators between normal and brain damaged school age

children.

Relatively little research has been reported with regard to the predictive validity of either the WISC or the Bender when used with very young school age children. While clinicians have gathered support for the use of both tests for concurrent diagnosis (Maling, 1973), the value of the tests as predictors of future test performance remains unclear. This research issue is of particular importance in view of the economic realities of psychological testing in the public schools. Individual tests like the Bender or the WISC are costly and, therefore, rarely given annually to each child. Instead, the child's performance as tested early in his school year may continue to be used as a basis for later evaluation based upon the assumption that the test scores are reliable and valid. Thus, WISC or Bender test scores which are recorded in a student's cumulative test file may be many years old, but still in current use by teachers and administrators (Gredler, 1975).

In order to better understand what those score reports mean a year, or two years, or five years after they were first recorded, further investigation must be conducted. With respect to the intellectual level and academic performance of the LD child as it is reflected in such scores, longitudinal research investigating the relationships between early test score data and later school performance is necessary.

The purpose of this study was to investigate the predictive validity of the WISC and the Bender-Gestalt over a four year period when used with learning disabled, primary school children. The relationships between individual test scores and later reading and math grade equivalent scores, group IQ scores, and educational placement were examined.

Procedure

Data were collected in a suburban Connecticut community during the fall of 1970 and the spring and fall of 1971. Subjects in the initial sample were 225 elementary school children between the ages of six and seven. All subjects had been selected by their teachers to attend "pre-primary" (learning disabled) classes. Each child was individually examined by a school psychologist and was administered the WISC and the Bender. The subjects were clinically categorized in one of the following groups: intellectual deficit, emotional dysfunction, perceptual dysfunction, any combination of the prior groupings, and no apparent dysfunction but suspected maturational lag. Clinical categorizations of the children were made by two school psychologists.

From the original group of subjects, follow-up data were collected on a sample of 104 students during the spring of 1976. These data included Otis-Lennon IQ scores, teachers' reports as to the grade level equivalent performance of the students in the areas of reading and math, and the current placement of the child (either regular classroom only or regular classroom plus some type of specific supportive service).

A preliminary multivariate analysis of variance (MANOVA) was performed to determine if the follow-up group was different from the original pool of subjects. Individual test score data collected during the 1970-71 evaluations plus biographical information as to the child's age at the time of the evaluation and the child's sex were the dependent variables. Group membership--follow-up or attrition--was the independent variable.

The data were then examined by a series of stepwise multiple regression analyses. Predictor variables included: (1) scaled subtest and full test scores from the Bender and the WISC; (2) the difference between the subject's Bender score and the norm for that age group; (3) the

difference between the subject's Verbal and Performance test scores on the WISC; (4) sex; and (5) age in months at the time of the evaluation. Current reading and math grade equivalent scores, and Otis-Lennon IQ scores were used as the criteria in the three separate analyses.

In the final analysis of the data, a stepwise discriminant analysis technique (DFA) was used. In order to determine how well individual test score data could predict the later educational needs of the students, Bender and WISC scores were used to predict current educational placement. The DFA was employed to classify subjects into one of two groups, either regular classroom instruction only ($N = 67$) or regular classroom plus supportive services ($N = 37$).

Results and Conclusions

The preliminary MANOVA yielded a non-significant F-ratio, suggesting that there was no selective attrition operating over the five year period.

Prediction of later math and reading achievement. The multiple regression analyses to predict grade level equivalence scores for reading reached an optimum R of .18 ("optimum" defined as minimum standard error of estimate). At that step, three predictors were included in the regression equation: full scale WISC score, total Bender error score, and the Bender difference score (mean score for subject's age group minus subject's Bender score).

In predicting math achievement, a similar R of .18 was obtained. The students' age was the only variable included in the equation. Because both R's were non-significant, the relationship between all predictors and reading and math achievement was no better than chance alone.

The lack of predictive power of the early cognitive measures when applied to reading and math scores five years later supports the idea that

achievement in these academic areas may be relatively independent of early IQ and psychomotor assessment. It is quite possible that within the IQ range studied here, reading and math facility are more the result of school and/or home training than the result of measurable cognitive potential. If the cognitive potential merely sets the upper boundary on one's ability, then specific school experiences may be the factors to be investigated in predicting the extent to which potential is realized. This result, while non-significant statistically, may be interpreted as a very positive finding in support of school services such as a resource room, teacher aides for individual instruction, remedial instruction in math and reading, or any such program which is available for the student with potential for academic problems.

There is, however, another interpretation of the finding which is not so positive. Because of the lack of support for the predictive validity of early IQ scores when applied to actual classroom skills, teachers, counselors, administrators, and parents need to take care not to make inappropriate assumptions about students' abilities based upon old test data. If the assumption is made that a fifth grade student will not be able to do grade level math work because of a low IQ score recorded when the student was in first grade, the assumption itself may lead to depressed academic performance on the part of the student. The poor math performance demonstrated by a fifth grader may be due to a lack of cognitive ability first observed at the first grade level, but the results of the multiple regression analyses explored here do not support such an explanation. The experimenter bias effect (EBE) as proposed by Rosenthal (1973) might serve as a better explanation for the student's poor performance. A teacher's preconceived notion as to the potential of the students in his class, either a positive or negative expectation, can have an influence on the

students' actual performance. When the expectation is a positive one, as in the original Rosenthal and Jacobson study (1968), and the effect is also positive, there is little need to worry about student performance. But in the case where the expectations and the effects are negative, as in the case of the fifth grade student who scored poorly in the first grade, then there is need to worry about stunting student performance as the result of an EBE. This negative EBE is a potential threat to any students scoring below the average on standardized tests, particularly when the test scores are maintained for years, without review or update, in students' cumulative files.

Educators should face the question of whether or not information that is not clearly valid for use at a later date should be maintained in a student's file. If the information is not relevant to the prediction or understanding of a child's later classroom performance, then why should it follow him through his school career?

Prediction of later IQ scores. The results of the attempts to predict later IQ scores were much different from those predicting reading and math achievement scores. The optimum step in the regression equation to predict Otis-Lennon group IQ scores reached an R of .61. The subjects' age at the time of evaluation, total Bender score, Bender difference score, and the WISC verbal score were included in the equation.

The sizable increase in the R for the prediction of the Otis-Lennon scores when compared to the reading and math scores suggests an important difference in the nature of the criteria. The IQ scores, even when predicting group test scores from individual test scores over a five year period, appear to be much more stable and thus predictable over time. The relatively greater predictability of the IQ scores as compared with reading

and math achievement further supports the idea that academic skills may be primarily dependent on training and experience, while IQ scores are more related to innate abilities or at least to early childhood experiences (Bloom, 1964). These results also suggest that measures of general cognitive functioning, while they should not be interpreted as predictors of later classroom skills, are relatively stable from an early age and that early IQ estimates may be acceptable as evidence for requesting federal or state funds to support programs for exceptional students.

Prediction of special class inclusion. In the final analysis, DFA was used to examine group separation (between regular class, and regular class plus some special service). The stepwise DFA was terminated when the next predictor variable had an F-to-enter of less than 1.00; the multivariate F ratio was 4.13 ($p < .01$) at that step. The eight predictor variables to enter the discriminant equation were scores on the WISC digit span, comprehension, block design, information, and vocabulary subtests, the Bender scores for integration, rotation, and distortion. Eighty-six percent of the regular classroom group were correctly classified, fifty-six percent of the students receiving supplementary services were correctly classified. This difference in the power of classification between the two groups presents a problem for school personnel who wish to use early data for later educational placement decisions. The financial, educational, and emotional cost of inappropriately recommending regular class placement for the student who needs special educational services must be considered carefully. The greater percentage of error in the prediction of the actual special services placement group suggests that a more conservative identification procedure than the one explored here should be used in the schools. Although the DFA reached a significant

F level, and the group separation and classification was moderate, the outcomes imply that more current data are required to make valid educational predictions and decisions.

Yet the DFA does provide some information for administrative planning. Data collection at an early stage in a child's school career may be insufficient for making individual student decisions, but very valuable in estimating what the later needs of the school system will be. From the early collection of data describing exceptional students, long range educational needs for the school system as a whole may be predicted. Upper limits of the number of students in a school system requiring special services over a five year period may be estimated, thus allowing budget and personnel planning to begin at an earlier stage than is currently the case.

Bibliography

- Bender, L. A Visual-Motor Gestalt Test and Its Use. Research Monograph of the American Orthopsychiatric Association, 1938, No. 3.
- Bender, L. Use of the Visual-Motor Gestalt Test in the diagnosis of learning disabilities. Journal of Special Education, 1970, 4, 29-39.
- Bloom, B. Stability and Change in Human Characteristics. New York: Wiley, 1964.
- Clements, S. (Ed.) Minimal brain dysfunction in children. NINBD Monograph No. 3. Washington: U. S. Government Printing Office, 1966.
- Clements, S. and Peters, J. Minimal brain dysfunction in the school age child. Archives of General Psychiatry, 1962, 6, 3, 185-197.
- Eaves, L., Kendall, D. and Crichton, J. The early detection of minimal brain dysfunction. Journal of Learning Disabilities, 1973, 2, 8, 454-462.
- Graham, F. and Berman, P. Current status of behavioral tests for brain damage in infants and pre-school children. American Journal of Orthopsychiatry, 1961, 31, 713-727.
- Gredler, G. (Ed.) Ethical and Legal Factors in the Practice of School Psychology. Harrisburg, PA: Pennsylvania State Department of Education, 1975.
- Hanvik, L., Nelson, S., Hanson, H., Anderson, A., Dressler, W., and Zarling, V. Diagnosis of cerebral dysfunction in children. American Journal of Diseases in Children, 1961, 101, 364-375.
- Holroyd, J. When WISC verbal IQ is low. Journal of Clinical Psychology, 1968, 24, 457.
- Maling, B. Multiple discriminant analysis as a technique for predicting rationally determined categories of learning disability. Unpublished doctoral dissertation, University of Connecticut, 1973.
- McConnell, O. Kopppitz's Bender-Gestalt scores in relation to organic and emotional problems in children. Journal of Clinical Psychology, 1967, 23, 374-379.
- Rosenthal, R. The Pygmalion effect lives. Psychology Today, 1973, 7, 4, 56-63.
- Rosenthal, R. and Jacobson, L. Pygmalion in the Classroom. New York: Holt, 1968.
- Rourke, B. Brian-behavior relationships in children with learning disabilities. American Psychologists, 1975, 30, 9, 911-920.
- Schulberg, H. and Tolor, A. Use of the Bender-Gestalt in clinical practice. Journal of Projective Techniques, 1961, 25, 347-351.
- Wagner, E. and Murray, A. Bender-Gestalts of organic children: Accuracy of clinical judgments. Journal of Projective Techniques and Personality Assessment, 1969, 33, 3, 240-242.
- Wechsler, D. Wechsler Intelligence Scale for Children Test Manual. New York: Psychological Corp., 1949.
- Wechsler, D. Wechsler Intelligence Scale for Children-Revised Test Manual. New York: Psychological Corp., 1974.
- Weiner, G. The Bender-Gestalt as a predictor of minimal neurological deficit in children eight to ten years of age. Journal of Nervous and Mental Disease, 1966, 143, 3, 275-280.